Knowledge Map Creation for Modeling Learning Behaviors in Digital Learning Environments

AUTHOR(S):

Flanagan, Brendan; Majumdar, Rwitajit; Akçapnar, Gökhan; Wang, Jingyun; Ogata, Hiroaki

CITATION:

Flanagan, Brendan ...[et al]. Knowledge Map Creation for Modeling Learning Behaviors in Digital Learning Environments. Companion Proceedings of the 9th International Conference on Learning Analytics and Knowledge (LAK'19) 2019: 428-436

ISSUE DATE:

2019-03

URL:

http://hdl.handle.net/2433/243240

Knowledge Map Creation for Modeling Learning Behaviors in Digital Learning Environments

Brendan Flanagan[1*], Rwitajit Majumdar[1], Gökhan Akçapınar[1,2], Jingyun Wang[3] and Hiroaki Ogata[1]

[1] Kyoto University, Japan
[2] Hacettepe University, Turkey
[3] Kyushu University, Japan
[*] flanagan.brendanjohn.4n@kyoto-u.ac.jp

ABSTRACT: There has been much research that demonstrates the effectiveness of using ontology to support the construction of knowledge during the learning process. However, the widespread adoption in classrooms of such methods are impeded by the amount of time and effort that is required to create and maintain an ontology by a domain expert. In this paper, we propose a system that supports the creation, management and use of knowledge maps at a learning analytics infrastructure level, integrating with existing systems to provide modeling of learning behaviors based on knowledge structures. Preliminary evaluation of the proposed text mining method to automatically create knowledge maps from digital learning materials is also reported. The process helps retain links between the nodes of the knowledge map and the original learning materials, which is fundamental to the proposed system. Links from concept nodes to other digital learning systems, such as LMS and testing systems also enable users to monitor and access lecture and test items that are relevant to concepts shown in the knowledge map portal.

Keywords: Knowledge map; concept-based analytics; concept maps; knowledge extraction;

1 INTRODUCTION

It has been well documented that learners can benefit from the use of maps to represent the key concepts of knowledge (Lee et al., 2012). Ausubel (1963; 1968) defined the effective assimilation of new knowledge into an existing knowledge framework as the achievement of “meaningful learning”, by which knowledge maps can serve as a kind of scaffold to help learners to organize knowledge and structure their own knowledge framework (Novak et al., 2006). However, the process of creating and maintaining these maps often involves a domain expert manually creating the knowledge map based on their experience and previous knowledge (Wang et al., 2017).

To support the creation and use of knowledge maps by teachers and learners, we propose a knowledge map system that integrates with existing digital learning environments and learning analytics infrastructure. To assist in the creation of knowledge maps from digital learning materials, we propose a process for extracting key concepts from unstructured text to generate knowledge structures. Maps that have been generated are stored in a Knowledge Map Store (KMS) and an authoring system is provided for teachers to create, edit, and manage stored knowledge maps before publishing. The Knowledge Portal provides visualizations of knowledge maps with attributes determined from the analysis of learning behavior event log data from existing learning analytics infrastructure. In the final section of this paper, we outline the anticipated cases in which the system will be utilized by both teachers and students to monitor individual and group knowledge states.

There are many previous researches into the generation and use of ontologies, concept maps, and knowledge maps in education to show and create knowledge frameworks. Association rules and other data mining techniques have been used to construct concept maps based on the results of test and quizzes to show the relation between knowledge that was tested (Hwang, 2003; Tseng et al., 2007; Chen et al., 2010; Chen et al., 2013). While this technique is applicable to the structured format of tests, it is difficult to apply similar techniques to unstructured text that is contained in digital learning materials.

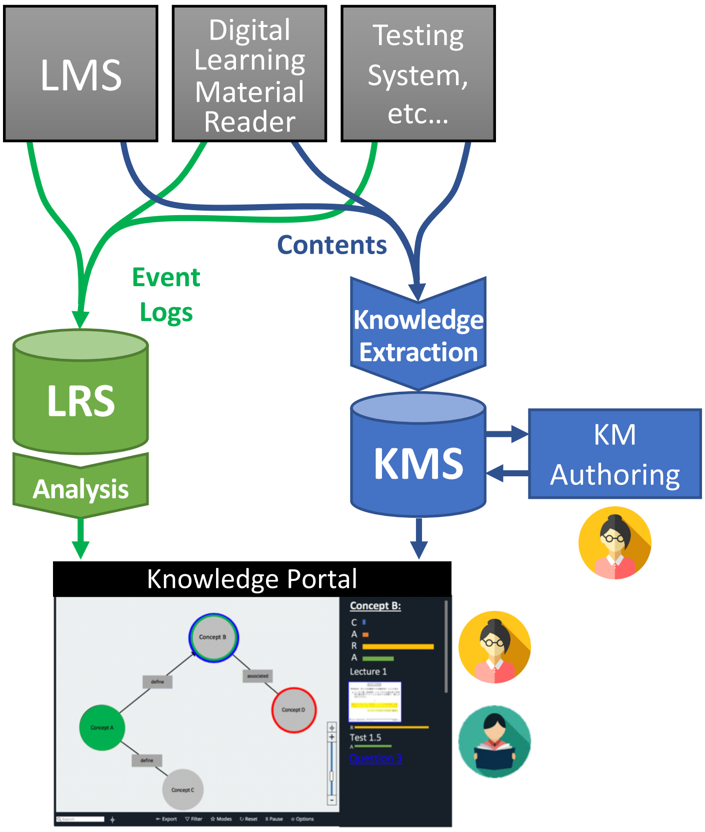

Figure 1: An overview of how the proposed system would integrate with existing LA infrastructure.

2 SYSTEM OVERVIEW

In this section, we provide an outline of the proposed Knowledge Map system, how it integrates with existing LA infrastructure, and how stakeholders will interact with the system. Fig. 1 shows an overview of the system with the main components consisting of:

- Existing user facing LA infrastructure, such as: LMS, Digital Learning Material Reader, Testing system, etc.
- LRS and Analytics Processor.
- Knowledge Extraction Processor.

- KMS (Knowledge Map Store) and a teacher facing Knowledge Map Authoring portal.
- User facing Knowledge Portal.

The existing user facing infrastructure, such as: LMS, Reader, and Testing system serve as an interaction event sensor and also as a source of learning material contents that are sent to the Knowledge Extraction Processor. Recent implementations of LA platforms often utilize an LRS and Analytics Processor as a pipeline for storing and processing event statement data about the use of user facing learning systems (Chatti et al., 2017; Flanagan and Ogata, 2017). We use this existing pipeline to provide information to augment the visualization of knowledge structures representing the underlying learning materials, lecture attendance, and past academic achievement.

Figure 2: Hierarchy of node attributes based on event log analysis.

The main hierarchy of node attributes based on analytics is shown in Fig. 2, where each level is linked to important stages in the formal learning process: lecture attendance, reading learning materials, confirming acquired knowledge through the answering of tests, and attaining a credit for having satisfied the requirements of a course. The most basic form of effort by a learner is to attend a lecture in which learning material related with the concept node was covered. When a learner actively reads the learning material the concept node is attributed as Read. If a learner has correctly answered a test item relating to the concept, then the node is given the Answered attribute. Finally, the if the student passes the course then the Credit attribute is assigned.

The Knowledge Extraction Processor analyzes learning content data from the LMS, Reader, and Testing system. In the present paper, we focus on the extraction of knowledge maps from PDF contents that have been uploaded to the digital learning material reader. The results of this process are then stored in the KMS. Teachers are able to manage knowledge maps stored in the KMS through a teacher facing authoring portal.

3 KNOWLEDGE MAP EXTRACTION FROM CONTENTS

Course curriculum in K-12 education is often well structured and defined by government level organizations that regulate education. However, higher education often is less regulated with the course curriculum being decided by the teacher. In Japanese universities, teachers in charge of courses are busy and course contents are often finished close to when a lecture is due to start, allowing little time to create knowledge maps manually.

The authoring section of the proposed system automatically analyzes contents uploaded by teachers to support the generation of knowledge maps. As a part of the authoring process, the system requires the map to be checked by the teacher before being used by students. The teacher is also able to edit the automatically generated knowledge map to add, remove, or alter required sections.

A knowledge map can be thought of as a graph of key points that are contained within the digital learning material contents that it represents. The relation between nodes of this graph are expressed as a weighted edge representing the strength of the relation between two key points that are in the contents. In this paper, we use a process based on a method previously proposed by Flanagan et al. (2013) as shown in Fig. 3.

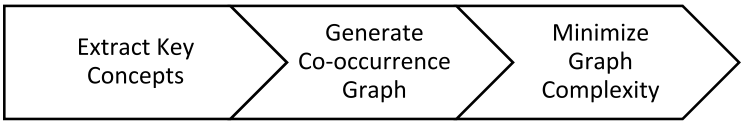

Figure 3: The process used to extract a knowledge map from digital learning material contents.

The lecture slides are usually written in Japanese with sections also in English. The text is extracted from lecture slides PDF files using pdfminer[1] and parsed with MeCab (Kudo, 2006) to separate individual words and parts-of-speech (POS) from a sentence using morphological analysis. Key concept terms are extracted by selecting the longest sequences of nouns and conjugate particles in a sentence. These were then indexed using the GETAssoc[2] search engine to form a co-occurrence matrix of terms. The link between the concept terms and the sections of the learning material are also included as an attribute in the search engine so relevant learning resources can be retrieved. The final step of the process involves minimizing the complexity of the co-occurrence graph using a minimum spanning tree algorithm to select the strongest concept term relations. In this implementation a thesaurus of technical terms in Japanese and English was used to guide the knowledge map generation process with hierarchical selection.

Table 1: Learning materials for the evaluation.

Lecture	Pages	Concepts (Gold Standard)	Max Concepts (Proposed)
1	30	12	125
2	32	10	153
3	45	6	222

We conducted a preliminary experiment using the proposed method in a university course on Information Science. A knowledge map that includes the concepts of three lecture learning materials was created manually by the course teacher and used as the gold standard for evaluation as shown in Table 1. Knowledge maps were automatically generated for each lecture with the strongest relation calculated using the SMART weight as described in Salton (1983). The precision/recall evaluation when comparing generated maps to the gold standard is shown in Fig. 4 with maximum precision of 0.72 at

1 https://euske.github.io/pdfminer/

2 http://getassoc.cs.nii.ac.jp

a threshold of 11 nodes for each generated map. As the threshold is increased the precision decreases, however the evaluation shows a majority of correct nodes are extracted at low thresholds. The generated knowledge maps would require some manual editing by a teacher before use in order to represent the same structure as the gold standard, and therefore is an ongoing topic of research.

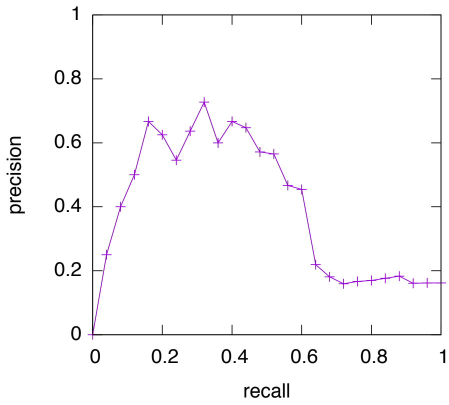

Figure 4: Plot of precision recall of proposed method

4 KNOWLEDGE MAP STORE AND AUTHORING

A centralized storage system of curated knowledge maps is fundamental to the analysis of knowledge accumulated over long-time spans. At the center of the proposed system, a KMS (Knowledge Map Store) acts as an LRS would for a conventional LA platform, collecting data about learning materials from disparate tools and systems to reduce information silos. This could enable the cross referencing and merging of knowledge maps from separate courses, learning materials, and even educational institutions if a KMS is deployed at the inter-institutional level.

The key data that a KMS should store are:

- The structure of knowledge maps that have been generated automatically by the system or created manually by teachers using the authoring interface.
- Links from the concept nodes of a knowledge map to related lecture schedule, learning materials, test items, and learner academic achievement records.

We are proposing that the structure of the knowledge map and links to learning materials/test items should be stored using a standards-based RDF storage service.

The proposed system has an authoring portal to facilitate the creation and management of knowledge maps by teachers. Automatically generated maps are initially stored as a draft and are not publicly available until the course teacher has confirmed the structure and its link to learning materials/test items. Maps can be edited to remove irrelevant nodes and add nodes that are required to cover the concepts in the course. A search function similar to the proposed knowledge map extraction process can be used to support the linking of relevant sections of learning materials and tests items to manually added nodes.

Knowledge maps can also be related with global concepts in the KMS to support large scale knowledge mapping across multiple courses. This feature is intended to facilitate the analysis of prior learner knowledge, thus allowing a teacher or learner to view what concepts learners have and have not acquired. There is also potential to apply the results of the analysis to recommend learning materials that should be studied to fill in knowledge gaps before attending a course.

5 KNOWLEDGE PORTAL

Once a knowledge map has been published with the authoring tool, it is available for use by students and teachers in the Knowledge Portal. The visualization interface for the proposed Knowledge Portal is based on a web-based open source ontology visualization system called WebVOWL (Lohmann et al., 2014). The interface of the proposed system is shown in Fig. 4 with the main knowledge map visualization on the left, and the right frame displays detailed information about the attributes of the selected node with relation to relevant learning materials. At the top of the right frame the use is given an overview of the percentile rank for each of the attributes: attend, read, and answer. It will also show if a credit has previously been attained in relation to the node concept. The user is able to follow the links to study learning materials or confirm their knowledge by a test item on the node concept. The visualization also features a filter to select specific nodes/relations and reduce the complexity of the knowledge map using varying degrees of edge collapse.

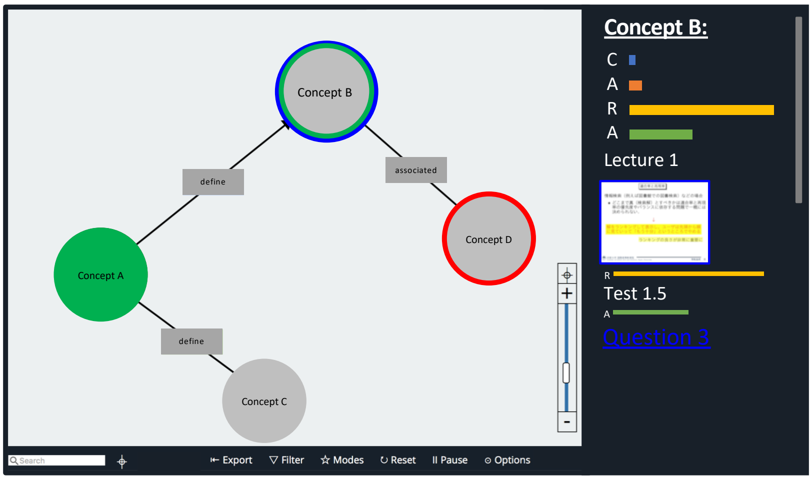

Figure 4: The user interface of the proposed system.

Additional functionality supports the augmentation of the base map structure with analytics results as visual attributes of nodes as was shown in Fig 3, to give uses visual ques to the overall knowledge state.

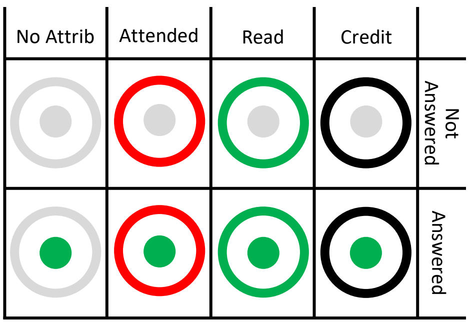

Figure 5: Node visual augmentation definition.

The outline color of a node represents the level of a learner's effort with the relevant learning materials that describe the node concept, and the fill color of the node relates to the learner's knowledge level of the node concept as shown in Fig 5. The degree of coloring in both the outline and fill are displayed to represent the percentile rank of achievement when compared to the whole student cohort. If there is no or very low percentile rank of event data for Attendance/Read/Credit and Answer relating to a node concept, then the outline and fill are displayed as grey.

6 USES OF THE KNOWLEDGE PORTAL

The following section outlines different cases in which the knowledge portal could be used to guide both teaching and learning. An overview of the four main cases is shown in Fig. 6.

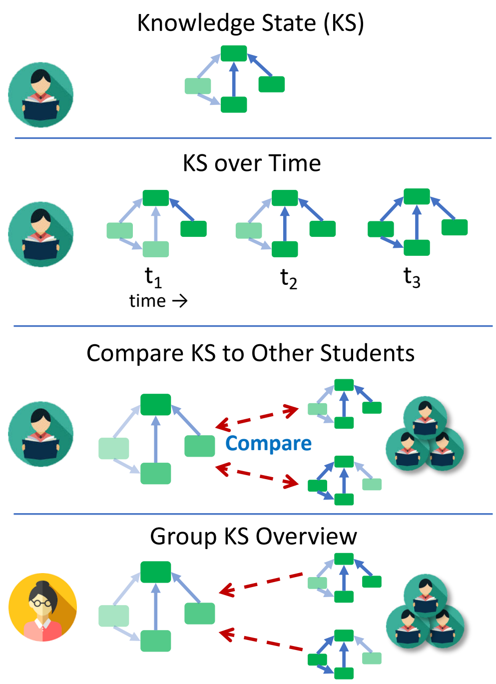

Figure 6: Different cases in which Knowledge Graphs could inform learns and teachers.

The first case is that of a learner confirming their current knowledge state for the support of self-regulated learning. It is intended that the learner could use the knowledge portal for the monitoring and planning of their learning by searching for concepts that they have not yet studied and following the links to appropriate learning materials to reading and test items to confirm their knowledge.

The second case enables the learner to reflect on how their knowledge has evolved over a period of time short or long, such as a student's knowledge at: t_1 = elementary school, t_2 = high school, t_3 = undergraduate university. This could also be used to help students find possible gaps in their knowledge that occurred in the past, and enable the revision of learning materials to resolve knowledge gaps.

The final two cases deal with comparing the knowledge state of groups of learners. For a student, this can enable them to compare their own knowledge to that of the broader student cohort and find possible areas in which their knowledge is lacking. The learner can then study to improve their knowledge state by working on specific concepts by reading learning materials and testing themselves with linked resources.

Teachers can also benefit from using the proposed knowledge map system to get an overview of the current knowledge state of all of the students in their course. The individual knowledge maps of all of the students are merged into a single aggregated knowledge map. An example use of this would be to check the prior knowledge of students before they attend a lecture, or checking the degree to which students have previewed concepts and the related learning materials to an upcoming class. The teacher then can adjust the lecture to either skip concepts that have been adequately learnt, or focus on concepts that require revision or greater explanation. It is expected that this case will be of particular use when managing courses with large numbers of students.

At a global knowledge map level, the relation between courses could provide insight into what parts of the knowledge map are important and central knowledge to a subject, and highlight what parts are difficult for students to understand and could be incorporated as a filter feature in the knowledge portal. This can be utilized in two different ways: for teachers it gives them an understanding of what knowledge is difficult to understand and may require more thorough explanation, and for students it allows them to see the knowledge that is central to the course and what areas they should pay attention to as it has been difficult for past students.

Knowledge map analysis could also be used in the recommendation of contents both inside and outside the course to learners based on their achievement and focus. Under achieving students may benefit from the recommendation of learning materials that cover concepts that they have yet to master. On the other hand, outperforming students may be interested in exploring extra learning materials outside of the course to expand their knowledge beyond that which would be traditionally offered.

7 CONCLUSIONS

In this paper, we proposed a system to support the creation, management and use of knowledge maps in digital learning environments at a learning analytics infrastructure level. In particular, we proposed processes for the automatic extraction, authoring, storing and use of knowledge maps by students and teachers. For the automatic extraction process, we proposed a text mining method for generating knowledge maps from digital learning materials and conducted a preliminary experiment to evaluate its effectiveness. A key feature of the method is the ability to link extracted concept nodes directly to specific parts of the learning materials from which they were extracted. These links are used to provide not only a reference for users to the original materials, but also as a method of associating learning behavior logs collected in existing system and mapping the analysis of these logs directly onto the knowledge map. This provides feedback to the user about the current learning behavior state overlaid on a knowledge structure.

In future work, the use of the knowledge portal to increase learner knowledge awareness and group formation by knowledge map clustering should be investigated.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Number 16H06304.

REFERENCES

Ausubel, D. P. (1963). The psychology of meaningful verbal learning, New York: Grune and Stratton.

Ausubel, D. P., Novak, J. D., & Hanesian, H. (1968). Educational psychology: A cognitive view (Vol. 6). New York: Holt, Rinehart and Winston.

Chatti, M. A., Muslim, A., & Schroeder, U. (2017). Toward an open learning analytics ecosystem. In *Big data and learning analytics in higher education* (pp. 195-219). Springer, Cham.

Chen, S. M., & Sue, P. J. (2013). Constructing concept maps for adaptive learning systems based on data mining techniques. *Expert Systems with Applications*, 40(7), 2746-2755.

Chen, S. M., & Bai, S. M. (2010). Using data mining techniques to automatically construct concept maps for adaptive learning systems. *Expert Systems with Applications*, 37(6), 4496-4503.

Flanagan, B., Yin, C., Inokuchi, Y., & Hirokawa, S. (2013). Supporting interpersonal communication using mind maps. *The Journal of Information and Systems in Education*, 12(1), 13-18.

Flanagan, B., & Ogata, H. (2017). Integration of Learning Analytics Research and Production Systems While Protecting Privacy, *Proceedings of the 25th International Conference on Computers in Education (ICCE2017)*, 333-338.

Hwang, G. J. (2003). A conceptual map model for developing intelligent tutoring systems. *Computers & Education*, 40(3), 217-235.

Kudo, T. (2006). Mecab: Yet another part-of-speech and morphological analyzer. http://taku910.github.io/mecab/.

Lee, J. H., & Segev, A. (2012). Knowledge maps for e-learning. *Computers & Education*, 59(2), 353-364.

Lohmann, S., Link, V., Marbach, E., & Negru, S. (2014). WebVOWL: Web-based visualization of ontologies. In *International Conference on Knowledge Engineering and Knowledge Management* (pp. 154-158). Springer, Cham.

Novak, J. D., & Cañas, A. J. (2006). The theory underlying concept maps and how to construct them. *Technical report IHMC CmapTools 2006-01 Rev 01-2008*. Florida Institute for Human and Machine Cognition.

Salton, G., & McGill, M. (1983). Introduction to modern information retrieval. McGraw-Hill, New York.

Tseng, S. S., Sue, P. C., Su, J. M., Weng, J. F., & Tsai, W. N. (2007). A new approach for constructing the concept map. *Computers & Education*, 49(3), 691-707.

Wang, J., Flanagan, B., & Ogata, H. (2017). Semi-automatic construction of ontology based on data mining technique. In *6th IIAI International Congress on Advanced Applied Informatics (IIAI-AAI)* (pp. 511-515). IEEE.